

# Score-Based Secretary Problem

Jyotirmoy Sarkar\*

Indiana University-Purdue University Indianapolis, USA  
jsarkar@iupui.edu

**Abstract.** In the celebrated “Secretary Problem,” involving  $n$  candidates who have applied for a single vacant secretarial position, the employer interviews them one by one in random order, and learns their relative ranks. As soon as each interview is over, the employer must either hire the candidate (and stop the process) or reject her (never to be recalled). We consider a variation of this problem where the employer also learns the scores of the already interviewed candidates, which are assumed to be independent and drawn from a known continuous probability distribution. With this additional information, what strategy should the employer follow in order to maximize his chance of hiring the candidate with the highest score among all  $n$  candidates? What is the maximum probability of hiring the best candidate?

**Keywords:** Analytical expression, Conditional probability, Iterative computation, Recursive relation, Simulation

## 1 Background and Statement of the Problem

Recall the celebrated Secretary Problem: There are  $n$  applicants who have applied for the single open position of a secretary. The employer will interview them one by one in random order. At the conclusion of each interview, the employer must decide either to hire the candidate (and stop interviewing) or to let her go for good never to be recalled. At any time the employer has a relative ranking of all candidates interviewed so far, but he does not have their absolute ranking among all  $n$  candidates (unless and until he interviews all candidates). His objective is to hire the best candidate. Obviously, he must consider hiring a candidate only if she is the best among all interviewed candidates. How can he maximize his probability of hiring the best candidate among all  $n$  candidates?

If the employer hires a candidate after interviewing too few candidates, he runs the risk of missing the best candidate who is yet to be interviewed. On the other hand, if he waits too long to hire, he will have too few candidates left to be interviewed, and hence a small chance that one of them will surpass the leading candidate among those who have been already interviewed (and this leading candidate has not been hired). Assuming that all permutations of candidates

---

\* The author thanks the Indian Statistical Institute Kolkata and Calcutta University Department of Statistics for hosting his sabbatical leave visit, while this research was conducted.

by rank are equally likely, the optimal strategy turns out to be: “Interview and let go the first  $m - 1$  candidates, and thereafter hire the first candidate among  $m, m + 1, \dots, n - 1$  who is the best among all interviewed candidates, if such a candidate comes along; otherwise, hire the last candidate.” The value  $m$ , of course, depends on  $n$ . In fact,  $m = \operatorname{argmax}_k \{p_n(k)\}$  where  $p_n(1) = 1/n$  and for  $k \geq 2$ ,  $p_n(k) = \frac{k-1}{n} \sum_{j=k}^n \frac{1}{j-1}$ . The maximum probability of hiring the best candidate by following the optimal strategy is  $p_n(m)$ . In particular, asymptotically (as  $n \rightarrow \infty$ ), the optimal choice is  $m \approx n/e$ , and the associated highest probability of hiring the best candidate is  $1/e = 0.3678794 \dots$ .

The secretary problem is also known as the marriage problem, the sultan’s dowry problem, the fussy suitor problem, the game of googol and the best choice problem. It was apparently introduced in 1949 by Merrill M. Flood [4], who mentions that R. Palermo proved that all strategies are dominated by a strategy of the form: “Reject the first  $k-1$  unconditionally, then accept the next candidate who is the best.” The first publication was apparently by Martin Gardner in *The Scientific American*, February 1960, where he presented the analysis by Leo Moser and J.R. Pounder. It was reprinted with additional comments in [5]. The  $1/e$ -law of best choice is due to F. Thomas Bruss [2]. We refer the reader to Ferguson [3] for an extensive bibliography.

Here we consider a variation of the Secretary Problem, in which we know the probability distribution from which the scores of each candidate is drawn. Suppose that the candidates appear in a random order. After interviewing each candidate, the employer not only knows her relative rank among all candidates interviewed so far, but also he knows her absolute score, which is assumed to be drawn from a known distribution. In this modified situation, what strategy should the employer follow in order to maximize his chance of hiring the candidate with the highest score among all  $n$  candidates? More precisely, we consider the problem below.

**Score-Based Secretary Problem:** In the context of the Secretary Problem, suppose that the employer, after interviewing the candidates, can assign absolute scores  $X_1, X_2, \dots$ , which are assumed to be drawn independently from the same known continuous distribution function  $F$ . What is the employer’s best strategy to maximize the chance of winning (that is, hiring the candidate with the highest score among all  $n$  candidates)? What is the maximum probability of winning?

In preparation for solving the score-based secretary problem, let us make a few straight-forward observations:

- 1) Without loss of generality, we can assume that the scores are independent and identically distributed (IID) as uniform  $(0, 1)$ . For otherwise, we will simply replace each score  $X$  by  $U = F(X)$ , which will have uniform  $(0, 1)$  distribution. See Exercise 1.2 of [7]. Also, by the continuity of  $F$ , ties among scores is ruled out (for when the scores are displayed to infinitely-many decimal places, surely they will differ).

- 2) The employer will surely let go any candidate who scores below someone else already interviewed. He should consider hiring a candidate only if her score is a **record high** value in the sense that it is the largest among all scores assigned so far. Clearly,  $X_1$  is a record high score. Thereafter, for  $m > 1$ ,  $X_m$  is a record high value if  $X_m = \max\{X_1, X_2, \dots, X_m\}$ ; or equivalently, if and only if  $X_m > X_i$  for  $i = 1, 2, \dots, m-1$ .
- 3) The employer should consider hiring a candidate with a record high score provided her score is large enough so that there is only a small chance that some other candidate yet to be interviewed will score higher than the current candidate. How large a score is large enough for the employer to win the game (or hire a candidate with record high score)?
- 4) The requisite threshold, above which the employer should hire a candidate with a record high score, depends on how many candidates are yet to be interviewed. Finding the threshold corresponding to each possible remaining number of candidates to be interviewed is the crux of the solution to the problem. We contend that the thresholds can be found inductively as we allow more and more candidates.

Let us describe the employer's best strategy and his overall chance of winning.

**Definition 1** If there are  $n$  applicants, let  $\theta_n$  denote the threshold such that the employer maximizes his probability of hiring the best candidate by using the strategy "Hire Candidate 1, who scores  $X_1$ , if and only if  $X_1 > \theta_n$ ."

**Form of the Best Strategy:** "If  $X_1 > \theta_n$ , then hire Candidate 1. Otherwise, wait until a candidate receives a record high score  $X_m$ . If  $X_m > \theta_{n+1-m}$ , then hire Candidate  $m$ . Otherwise, let her go and wait until the next record high score is attained. Etc."

**Definition 2** If there are  $n$  applicants, let  $\omega_n$  denote the (maximal) probability that the employer wins the game (that is, hires the best candidate) if he uses his above-stated best strategy.

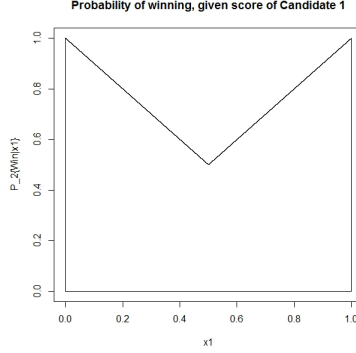
We will now inductively determine  $\theta_n$  and  $\omega_n$  for all  $n = 1, 2, 3, \dots$

## 2 Exact Solutions for Small Number of Candidates

Recall that the potential scores of the candidates (in the order interviewed) are  $X_1, X_2, \dots, X_n$ , which we assume are IID uniform  $(0,1)$ . If there is only  $n = 1$  applicant, the solution is trivial: Hire her irrespective of her score; that is,  $\theta_1 = 0$ . In this case the employer surely hires the best candidate by default. So,  $\omega_1 = 1$ . Let us proceed to determine  $(\theta_2, \omega_2), (\theta_3, \omega_3), \dots$

Suppose that  $n = 2$ . In this case, after interviewing Candidate 1 who has received score  $X_1$ , there is only one more candidate to be interviewed. If  $x_1 > 1/2$ , then the employer should hire Candidate 1, because, given  $x_1 > 1/2$ , the conditional probability that the employer wins (or hires the better candidate) is  $P\{X_2 < x_1\} = x_1 > 1/2$ . On the other hand, if  $x_1 < 1/2$ , then the employer should let Candidate 1 go, and hire Candidate 2, because, given  $x_1 < 1/2$ , the conditional probability that the employer wins is  $P\{X_2 > x_1\} = 1 - x_1 > 1/2$ .

Fig. 1 shows the function  $f_2(x_1) = \max\{1 - x_1, x_1\}$ , the conditional probability of the employer winning, given  $X_1 = x_1$ .



**Fig. 1.** Conditional probability  $f_2(x_1)$  that the employer wins, given  $X_1 = x_1$

Since  $X_1$  follows a uniform  $(0, 1)$  distribution, the overall probability that the employer wins is the area under the graph of  $f_2(x_1)$ ; that is,

$$\omega_2 = \int_0^1 f_2(x_1) dx_1 = \int_0^{1/2} (1 - x_1) dx_1 + \int_{1/2}^1 x_1 dx_1 = \frac{3}{4}. \quad (1)$$

We claim that  $\theta_2 = 1/2$ . The justification is as follows: If the employer adopts a strategy of hiring Candidate 1 if and only if  $x_1 > \theta$ , with any other threshold value  $\theta$ , then his overall probability of winning will be, as seen from Fig. 2,

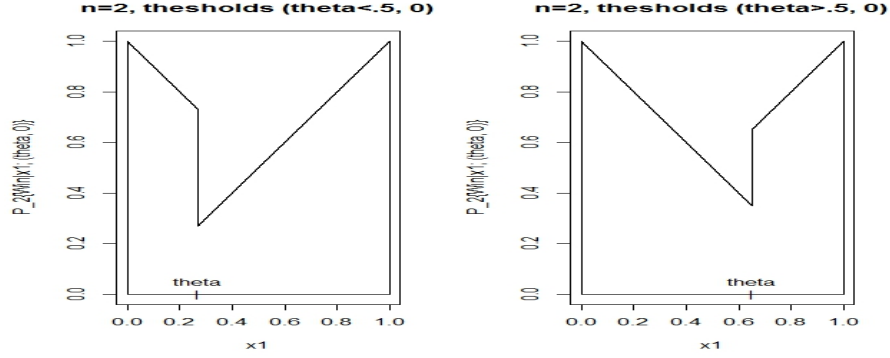
$$\int_0^\theta (1 - x_1) dx_1 + \int_\theta^1 x_1 dx_1 = \theta(1 - \theta) + 1/2 = \frac{3}{4} - \left(\theta - \frac{1}{2}\right)^2 \in \left(\frac{1}{2}, \frac{3}{4}\right). \quad (2)$$

Hence, any threshold value  $\theta$  other than  $\theta_2 = 1/2$  actually lowers the employer's overall probability of winning. Interestingly though, by choosing any sub-optimal threshold  $\theta_2 \neq 1/2$ , even by choosing  $\theta = 0$  (that is, always hire Candidate 1 irrespective of her score) or  $\theta = 1$  (that is, never hire Candidate 1), the employer's overall chance of winning never falls below  $1/2$ .

Next, suppose that  $n = 3$ . In this case, after interviewing Candidate 1, who scored  $X_1 = x_1$ , there are two more candidates to interview. We must determine the value of  $\theta_3$  such that the employer's best strategy is to hire Candidate 1 with score  $X_1 = x_1$  if and only if  $x_1 > \theta_3$ . Let us analyze separately what happens if (A) Candidate 1 is hired, and (B) Candidate 1 is let go:

(A) If Candidate 1 is hired, then the interviewer has not seen scores  $X_2$  and  $X_3$ , which are independent uniform  $(0, 1)$  variables. In this case, the employer's conditional probability of winning is

$$P_3^{\text{hire}}(x_1) = \Pr\{X_2 < x_1, X_3 < x_1\} = \Pr\{X_2 < x_1\} \cdot \Pr\{X_3 < x_1\} = x_1^2. \quad (3)$$



**Fig. 2.** For  $n = 2$ , the employer's probability of hiring the better candidate if he uses a suboptimal strategy by choosing thresholds  $\theta$  and 0 for Candidates 1 and 2 respectively

Note that  $P_3^{\text{hire}}(x_1)$  is an increasing function of  $x_1$  on  $(0, 1)$  going from 0 to 1. Thus, whenever it is advantageous for the employer to hire Candidate 1 with a particular score  $x_1$ , it must be also advantageous for him to hire Candidate 1 with any score bigger than  $x_1$ .

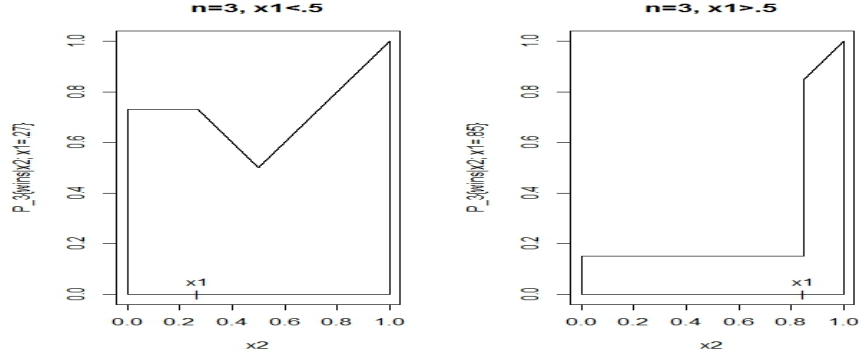
(B) On the other hand, if the employer lets go Candidate 1 with a score  $x_1$ , he gets to interview Candidate 2 and obtain  $X_2$ . The problem almost reduces to the case of only two candidates, except that the employer should not apply his best strategy for the two-candidate game unless  $X_2$  is a record high value; that is, unless  $X_2 > x_1$ . There are three mutually exclusive cases to consider: (1) If  $X_2 < x_1$ , then the employer should let Candidate 2 go, he should continue to interview Candidate 3 to obtain  $X_3$ , and win only if  $X_3 > x_1$ . (2) If  $X_2 > x_1$  but  $X_2 < \theta_2$ , then (in accordance with the optimal strategy for the two-candidate game) the employer should still let Candidate 2 go, interview Candidate 3 to obtain  $X_3$ , and win only if  $X_3 > X_2$ . (3) If  $X_2 > x_1$  and  $X_2 > \theta_2$ , then the employer should hire Candidate 2, and win if  $X_3 < X_2$ .

So, given  $x_1$  and the contemplated choice that the employer should let go (or skip over) Candidate 1, the conditional probability of his winning is

$$\begin{aligned}
 P_3^{\text{skip}}(x_1) &= Pr(\{X_2 < x_1, X_3 > x_1\} \cup \{X_2 > x_1, X_2 < \theta_2, X_3 > X_2\} \cup \\
 &\quad \{X_2 > x_1, X_2 > \theta_2, X_3 < X_2\}) \\
 &= x_1(1 - x_1) + \int_{x_1}^1 f_2(x) dx = \begin{cases} \frac{3}{4} - \frac{x_1^2}{2} & \text{if } x_1 \leq 1/2 \\ (1 - x_1)(x_1 + \frac{1+x_1}{2}) & \text{if } x_1 \geq 1/2 \end{cases} .
 \end{aligned}$$

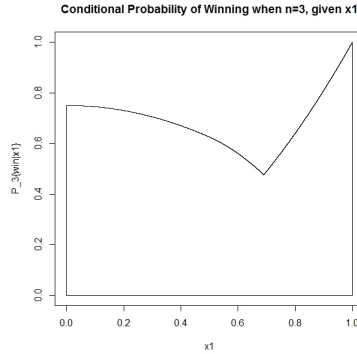
Fig. 3a-b show  $P_3^{\text{skip}}(x_1)$  as the total area of the enclosed region for two typical values of  $x_1$ —(a) below  $1/2$  and (b) above  $1/2$ .

It can be checked that  $P_3^{\text{skip}}(x_1)$  is a decreasing function of  $x_1$  on  $(0, 1)$  going from  $\omega_2 = 3/4$  to 0, while  $P_3^{\text{hire}}(x_1)$  is an increasing function of  $x_1$  on  $(0, 1)$  going from 0 to 1. By the intermediate value theorem (see [1], for example) these two



**Fig. 3.** For (a)  $x_1 < .5$  and (b)  $x_1 > .5$ , the conditional probability that the employer wins if he lets go Candidate 1, obtains  $X_2 = x_2$ , and follows his best strategy.

functions intersect in  $(0, 1)$  at a unique point  $x_1 = \theta_3 = (1 + \sqrt{6})/5$ , which is obtained by solving  $(1 - x_1) \cdot (1 + 3x_1)/2 = x_1^2$ , or equivalently  $5x_1^2 - 2x_1 - 1 = 0$ . Therefore, the employer should hire Candidate 1 if and only if  $x_1 > \theta_3$ . Hence, given  $x_1$ , the conditional probability of the employer winning (if he follows the optimal strategy) is  $f_3(x_1) = \max\{P_3^{\text{skip}}(x_1), P_3^{\text{hire}}(x_1)\}$ , as shown in Fig. 4.



**Fig. 4.** Conditional probability  $f_3(x_1)$  that the employer wins the three-candidate game, given Candidate 1's score  $x_1$ , if the employer follows his optimal strategy

Finally, if the employer uses this optimal strategy for  $n = 3$ , his overall probability of winning is

$$\omega_3 = \int_0^1 f_3(x) dx$$

$$\begin{aligned}
&= \int_0^{1/2} \left( \frac{3}{4} - \frac{x_1^2}{2} \right) dx_1 + \int_{1/2}^{\theta_3} (1 - x_1) \left( \frac{1 + 3x_1}{2} \right) dx_1 + \int_{\theta_3}^1 x_1^2 dx_1 \\
&= \frac{17}{48} + \frac{x_1 + x_1^2 - x_1^3}{2} \Big|_{1/2}^{\theta_3} + \frac{x_1^3}{3} \Big|_{\theta_3}^1 = \frac{293 + 48\sqrt{6}}{600} = 0.68429251.
\end{aligned}$$

For  $n \geq 4$  candidates, we proceed inductively. Suppose that the solutions for the  $n$ -candidate game— $P_n^{\text{hire}}(x_1)$ ,  $P_n^{\text{skip}}(x_1)$ ,  $f_n(x_1)$ ,  $\theta_n$ ,  $\omega_n$ —are already found. Then we find the solutions for the  $(n+1)$ -candidate game using Algorithm 1.

### Algorithm 1

- Step 1** Define  $P_{n+1}^{\text{hire}}(x_1) = x_1 \cdot P_n^{\text{hire}}(x_1) = x_1^n$ , which is an increasing function of  $x_1$  on  $(0, 1)$  going from 0 to 1.
- Step 2** Define  $P_{n+1}^{\text{skip}}(x_1) = x_1 \cdot P_n^{\text{skip}}(x_1) + \int_{x_1}^1 f_n(x_1) dx_1$ , which is a decreasing function on  $(0, 1)$  going from  $\omega_n$  to 0.
- Step 3** The unique solution to  $P_{n+1}^{\text{skip}}(x_1) = P_{n+1}^{\text{hire}}(x_1)$  in  $x_1$  yields  $\theta_{n+1}$ .
- Step 4** Define  $f_{n+1}(x_1) = \max\{P_{n+1}^{\text{skip}}(x_1), P_{n+1}^{\text{hire}}(x_1)\}$ , and obtain the winning probability  $\omega_{n+1} = \int_0^1 f_{n+1}(x_1) dx_{n+1}$ .

Using Algorithm 1, we can obtain  $\theta_4 = .775845$ , by solving  $17x_1^3 - 6x_1^2 - 3x_1 - 2 = 0$ ; and we can evaluate  $\omega_4 = \int_0^1 f_4(x_1) dx_1 = .655397$ . In this manner, we can continue to use Algorithm 1 to find  $(\theta_5, \omega_5), (\theta_6, \omega_6), \dots$ . But the Calculus becomes tedious! In Section 3 we describe a computational technique to evaluate  $(\theta_n, \omega_n)$  recursively, and we verify the optimality of the solutions through simulation. In Section 4 we develop analytical expressions of  $(\theta_n, \omega_n)$  for all  $n \geq 3$ .

## 3 Computing $(\theta_n, \omega_n)$ and Verifying by Simulation

Now-a-days the ability to compute is at everyone's fingertip—thanks to the invention of personal computer and the advancement of computer languages. For instance, using the programming language (and statistical software) R, which anyone can download for free on to their computer, one can evaluate  $(\theta_n, \omega_n)$ .

### 3.1 Computational Evaluation of $(\theta_n, \omega_n)$

First, the integrals in Steps 2 and 4 of Algorithm 1 can be approximated by computing the Riemann sum. To do so, we recommend partitioning the interval  $(0, 1)$  into  $10^6$  equal sub-intervals of width  $10^{-6}$  each and evaluate the integrand at the middle point of each sub-interval. Then the integral is approximated by the mean of the functional values at the  $10^6$  midpoints. See [1], for example.

Second, because  $P_{n+1}^{\text{skip}}(x_1)$  is decreasing from  $\omega_n$  to 0, while  $P_{n+1}^{\text{hire}}(x_1)$  is increasing from 0 to 1 in the interval  $(0, 1)$ , the  $n$ -degree polynomial equation in

Step 3 of Algorithm 1 can be solved (approximately) by computing the largest argument at which  $P_{n+1}^{\text{skip}}(x_1) > P_{n+1}^{\text{hire}}(x_1)$ , plus half of  $10^{-6}$ . The R codes for inductively computing  $(\theta_{n+1}, \omega_{n+1})$  are given below.

### R Codes to Compute $(\theta_{n+1}, \omega_{n+1})$ Inductively

```
B=10^6; H=0.5/B
serial=seq(1:B);grid=serial/B-H
theta=c(0); win=c(1) # for n=1
# for n=2 define directly
hire=grid;skip=1-grid;
theta=c(theta, grid[max(which(skip>hire))]+H)
fun=pmax(skip, hire); win=c(win, mean(fun))
# for n>2 define recursively
for (i in 3:100){
  hire=hire*grid # Step 1
  skip=(cumsum(fun[B:1])[B:1]+(serial-1)*skip)/B # Step 2
  theta=c(theta, grid[max(which(skip>hire))]+H) # Step 3
  fun=pmax(skip, hire); win=c(win, mean(fun)) # Step 4
}
```

Table 1 gives  $(\theta_n, \omega_n)$  for  $n = 1, 2, 3, \dots, 100$  correct to six decimal places. Fig. 5a-d depict the conditional probability  $f_n(x_1)$  that the employer wins given  $x_1$ , if he follows his optimal strategy for  $n = 5, 10, 20, 100$  candidates.  $(\theta_{n+1}, \omega_{n+1})$

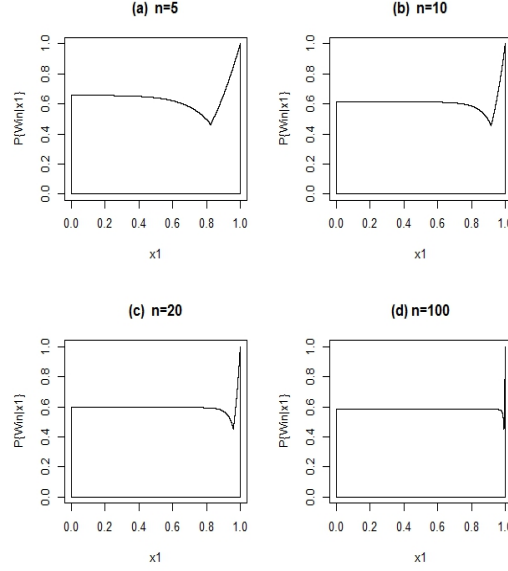
Table 1. The thresholds of the employer's optimal strategy and his maximal probability of winning corresponding to 1-100 candidates.

```
>round(theta, 6)
[1] 0.000000 0.500000 0.689898 0.775845 0.824590 0.855949 0.877807 0.893910 0.906265 0.916044
[11] 0.923976 0.930539 0.936059 0.940767 0.944829 0.948370 0.951484 0.954243 0.956706 0.958917
[21] 0.960913 0.962724 0.964375 0.965886 0.967274 0.968553 0.969736 0.970834 0.971854 0.972806
[31] 0.973695 0.974528 0.975310 0.976045 0.976738 0.977392 0.978010 0.978595 0.979150 0.979677
[41] 0.980178 0.980654 0.981109 0.981542 0.981956 0.982352 0.982731 0.983094 0.983442 0.983776
[51] 0.984097 0.984405 0.984702 0.984987 0.985262 0.985527 0.985783 0.986030 0.986269 0.986499
[61] 0.986722 0.986938 0.987146 0.987349 0.987544 0.987734 0.987919 0.988097 0.988271 0.988440
[71] 0.988603 0.988763 0.988917 0.989068 0.989215 0.989357 0.989496 0.989632 0.989764 0.989892
[81] 0.990018 0.990140 0.990260 0.990376 0.990490 0.990601 0.990710 0.990816 0.990919 0.991021
[91] 0.991120 0.991217 0.991312 0.991405 0.991495 0.991584 0.991672 0.991757 0.991841 0.991922

>round(win, 6)
[1] 1.000000 0.750000 0.684293 0.655396 0.639194 0.628784 0.621508 0.616128 0.611986 0.608700
[11] 0.606028 0.603813 0.601948 0.600356 0.598981 0.597781 0.596725 0.595788 0.594952 0.594201
[21] 0.593522 0.592906 0.592344 0.591830 0.591358 0.590922 0.590519 0.590145 0.589797 0.589472
[31] 0.589169 0.588885 0.588618 0.588367 0.588131 0.587908 0.587697 0.587497 0.587307 0.587127
[41] 0.586956 0.586793 0.586638 0.586490 0.586348 0.586213 0.586084 0.585960 0.585841 0.585726
[51] 0.585617 0.585511 0.585410 0.585312 0.585218 0.585127 0.585040 0.584955 0.584874 0.584795
[61] 0.584719 0.584645 0.584573 0.584504 0.584437 0.584372 0.584309 0.584248 0.584189 0.584131
[71] 0.584075 0.584020 0.583967 0.583916 0.583865 0.583817 0.583769 0.583723 0.583678 0.583633
[81] 0.583591 0.583549 0.583508 0.583468 0.583429 0.583391 0.583354 0.583317 0.583282 0.583247
[91] 0.583213 0.583180 0.583147 0.583116 0.583085 0.583054 0.583024 0.582995 0.582966 0.582938
```

As seen in Table 1, for the game involving 10 candidates, the employer's overall probability of winning, if he follows his optimal strategy, is about 61%. Likewise, for the game involving 100 candidates, the employer's overall probability of winning, if he follows his optimal strategy, is about 58%. However, we also





**Fig. 5.** Conditional probability  $f_n(x_1)$  that the employer wins, given the score  $x_1$  of Candidate 1, if he follows his optimal strategy for (a) 5, (b) 10, (c) 20, (d) 100 candidates

note that his probability of winning is monotonically decreasing as the number of candidates is increasing. Will the game continue to be favorable to him as the number of candidate increase, say to 500 or to 1000? Will there be a limiting value to the employer's probability of winning as the number of candidates increase without bound? Based on the fact that the thresholds and the winning probabilities for 951-1000 candidates have stabilized, as seen from Table 2, we conjecture that the answers are affirmative; but we leave the details to the reader to investigate.

Table 2. The thresholds of the employer's optimal strategy and his maximal probability of winning corresponding to 951-1000 candidates.

```
>round(theta[951:1000], 6)
[951] 0.999154 0.999155 0.999156 0.999157 0.999158 0.999158 0.999159 0.999160 0.999161 0.999162
[961] 0.999163 0.999164 0.999165 0.999165 0.999166 0.999167 0.999168 0.999169 0.999170 0.999171
[971] 0.999171 0.999172 0.999173 0.999174 0.999175 0.999176 0.999177 0.999177 0.999178 0.999179
[981] 0.999180 0.999181 0.999182 0.999182 0.999183 0.999184 0.999185 0.999186 0.999187 0.999187
[991] 0.999188 0.999189 0.999190 0.999191 0.999191 0.999192 0.999193 0.999194 0.999195 0.999195

> round(win[951:1000], 6)
[951] 0.580477 0.580477 0.580477 0.580476 0.580476 0.580476 0.580475 0.580475 0.580475 0.580475
[961] 0.580474 0.580474 0.580474 0.580473 0.580473 0.580473 0.580473 0.580472 0.580472 0.580472
[971] 0.580472 0.580471 0.580471 0.580471 0.580470 0.580470 0.580470 0.580470 0.580469 0.580469
[981] 0.580469 0.580469 0.580468 0.580468 0.580468 0.580468 0.580467 0.580467 0.580467 0.580467
[991] 0.580466 0.580466 0.580466 0.580466 0.580465 0.580465 0.580465 0.580464 0.580464 0.580464
```

### 3.2 Verifying Optimality of Computed Values via Simulation

In solving the Score-based Secretary Problem, we made claims of a probabilistic nature. For example, when  $n = 5$ , we claimed: “The employer’s maximal probability of winning a five-candidate game is  $\omega_5 = .639194$ , and he can achieve this probability by adhering to an optimal strategy with  $\theta_1 = 0, \theta_2 = .5, \theta_3 = .689898, \theta_4 = .775845, \theta_5 = .824590$ .” The correctness of this claim can be verified by conducting a simulation, which involves imitating a random phenomenon by generating an appropriate sequence of random numbers.

One can simulate the 5-candidate game a large number of times to estimate the employer’s probability of winning using the R codes below. The `win5` function, used in the codes, indicates whether the employer wins in any one particular play. The estimated probability of the employer’s winning,  $\hat{p}$ , is given by the mean of the indicators of win in a series of  $M = 10^6$  plays. In fact, in view of the central limit theorem, there is about 95% chance that the estimate  $\hat{p}$  will be within  $1.96\sqrt{p(1-p)/M} < 1/\sqrt{M}$  of the true probability  $p$ . See [2], for example. Therefore, our estimate  $\hat{p}$  based on  $M = 10^6$  repetitions, will be correct to three decimal places with about 95% probability.

When the above-described simulation is run with the optimal choice of the threshold values, the estimate of the employer’s probability of winning is found to be .638607, whereas when run with one particular alternative choice of thresholds the estimate turns out to be .636515, and with another alternative choice it turns out to be .592601. Since the estimates are correct to three decimal places, as reasoned above, this simulation provides sufficient justification for the optimal choice of thresholds, and declares the other choices of thresholds as suboptimal. We encourage the readers to estimate the probability that the employer wins when other choices of the threshold values  $\theta_3, \theta_4, \theta_5$  are made; and verify that none will do better than the optimal choices of thresholds our Algorithm 1 has produced. In Section 4 we will compute the employer’s exact probability of winning, for each of these strategies.

#### R Codes to Simulate Winning Probability for any Set of Thresholds

```
win5 <- function(theta3,theta4,theta5){
  coin=runif(5, 0, 1); top=max(coin)
  if (coin[1]>=theta5){ind=(coin[1]==top)}
  else{if (coin[2]>=max(coin[1],theta4)){ind=(coin[2]==top)}
    else{if (coin[3]>=max(coin[1:2],theta3)){ind=(coin[3]==top)}
      else{if (coin[4]>=max(coin[1:3],1/2)){ind=(coin[4]==top)}
        else{ind=(coin[5]==top)}
      }
    }
  }
  ind
}
mean(replicate(10^6, win5((1+sqrt(6))/5,.775845,.824590))) #opt
mean(replicate(10^6, win5(2/3, 3/4, 4/5))) # alternative 1
mean(replicate(10^6, win5(1/2,.5^(1/5),.5^(1/5),.5^(1/5)))) #2
```

#### 4 Analytical Expressions of $(\theta_n, \omega_n)$

A careful study of Algorithm 1 leads to an analytical expression for  $\theta_n$ , given in Theorem 1, and thereafter an analytical expression of  $\omega_n$ , given in Theorem 2. Towards these results, we note the following:

- (a) From Step 1, we have  $P_{n+1}^{\text{hire}}(x) = x^n$  for all  $x \in (0, 1)$ .
- (b) From Step 4, we have  $f_n(x) = \max\{P_n^{\text{skip}}(x), P_n^{\text{hire}}(x)\}$  for all  $x \in (0, 1)$ . In particular,  $f_n(x) = P_n^{\text{hire}}(x)$  for all  $x \in (\theta_n, 1)$ .
- (c) From Steps 2 and 1 and item (b) above, for all  $x \in (\theta_n, 1)$ , we have

$$P_{n+1}^{\text{skip}}(x) = x P_n^{\text{skip}}(x) + \int_x^1 P_n^{\text{hire}}(u) du = x P_n^{\text{skip}}(x) + \frac{1 - x^n}{n} \quad (4)$$

In fact, solving (4), we get Lemma 1 (to be proved momentarily):

**Lemma 1.** *For all  $n \geq 1$ , we have*

$$P_{n+1}^{\text{skip}}(x) = \left[ \sum_{k=1}^n \frac{x^{-k} - 1}{k} \right] x^n. \quad (5)$$

- (d) From Step 3, we have  $\theta_{n+1}$  is the unique solution to  $P_{n+1}^{\text{skip}}(x) = P_{n+1}^{\text{skip}}(x)$ ; or equivalently, in view of items (a) and (c) above,  $\theta_{n+1}$  is the solution to

$$\sum_{k=1}^n \frac{x^{-k} - 1}{k} = 1. \quad (6)$$

In other words, we obtain Theorem 1 (to be proved shortly):

**Theorem 1.**  $\theta_{n+1} = (1 + \epsilon_{n+1})^{-1}$ , where  $\epsilon_{n+1}$  is the unique solution to

$$\sum_{i=1}^n \binom{n}{i} \frac{\epsilon^i}{i} = 1. \quad (7)$$

**Proof of Lemma 1.** The proof is by mathematical induction on  $n$ . We know that  $P_1^{\text{skip}}(x) = 0$  and  $P_1^{\text{hire}}(x) = 1$  for all  $x \in (0, 1)$ . Next, we know that  $P_2^{\text{skip}}(x) = 1 - x$  for all  $x \in (0, 1)$ . Also, putting  $n = 1$  in (4), for  $x \in (0, 1)$ , we have

$$P_2^{\text{skip}}(x) = x P_1^{\text{skip}}(x) + \int_x^1 u^0 du = x \cdot 0 + (1 - x) = (x^{-1} - 1) x$$

Hence, (5) holds for  $n = 1$ .

Now suppose that (5) holds for an arbitrary  $n \geq 1$ . Then from (4), we have

$$\begin{aligned} P_{n+1}^{\text{skip}}(x) &= x \cdot \left[ \sum_{k=1}^{n-1} \frac{x^{-k} - 1}{k} \right] x^{n-1} + \frac{1 - x^n}{n} \\ &= \left[ \sum_{k=1}^{n-1} \frac{x^{-k} - 1}{k} \right] x^n + \frac{1 - x^n}{n} = \left[ \sum_{k=1}^n \frac{x^{-k} - 1}{k} \right] x^n. \end{aligned}$$

Hence, (5) holds for  $n + 1$ . This completes the proof of Lemma 1. Q.E.D.

**Proof of Theorem 1.** Write  $x = (1 + \epsilon)^{-1}$ . Then starting from (6), applying the binomial theorem, and simplifying, we have

$$\begin{aligned} 1 &= \sum_{k=1}^n \frac{x^{-k} - 1}{k} = \sum_{k=1}^n \frac{(1 + \epsilon)^k - 1}{k} \\ &= \sum_{k=1}^n \frac{1}{k} \sum_{i=1}^k \binom{k}{i} \epsilon^i = \sum_{i=1}^n \sum_{k=i}^n \frac{1}{i} \binom{k-1}{i-1} \epsilon^i \\ &= \sum_{i=1}^n \left\{ \sum_{k=i}^n \binom{k-1}{i-1} \right\} \frac{\epsilon^i}{i} = \sum_{i=1}^n \binom{n}{i} \frac{\epsilon^i}{i}. \end{aligned}$$

This completes the proof. Q.E.D.

We already know that  $\theta_1 = 0, \theta_2 = 1/2$ . Putting  $n = 2$  into Theorem 1, we solve  $2\epsilon + \epsilon^2/2 = 1$  to obtain  $\epsilon_3 = -2 + \sqrt{6}$ ; whence  $\theta_3 = (1 + \epsilon_3)^{-1} = (-1 + \sqrt{6})^{-1} = (1 + \sqrt{6})/5$ . Next, putting  $n = 3$  into Theorem 1, we solve  $3\epsilon + 3\epsilon^2/2 + \epsilon^3/3 = 1$  to obtain  $\epsilon_4 = 0.288917$ ; and hence  $\theta_4 = 1/1.288917 = .775845$ . The advantage of using Theorem 1, over the computational technique of Algorithm 1 in Section 3, is that we can now compute  $\theta_n$  directly without having to compute first  $\theta_3, \theta_4, \dots, \theta_{n-1}$ .

**Remark 1** Since the left hand side expression in (7) is increasing in  $n$  at each fixed  $\epsilon \in (0, 1)$ , the solution  $\{\epsilon_{n+1}\}$  is a decreasing sequence. Hence,  $\{\theta_n\}$  is an increasing sequence.

**Remark 2** For  $n \geq 100$ , we can approximate  $\theta_n$  as  $0.44765^{1/n}$ , or more simply approximate it as  $1 - 0.8035/n$ .

The optimal thresholds  $\theta_1 = 0, \theta_2 = 1/2, \theta_3 = (1 + \sqrt{6})/5, \theta_4, \dots, \theta_n$  are used to find an analytical expression for  $\omega_n$  as given in Theorem 2 below. In fact, this theorem holds for any non-decreasing sequence of thresholds  $\xi_1, \xi_2, \dots, \xi_n$ , which need not be the optimal thresholds. The proof is delegated to the Appendix.

**Theorem 2.** *For any non-decreasing sequence of thresholds  $\{\xi_n\}$  with  $\xi_n \in (0, 1)$ , used by the employer to hire a candidate whose score is a running maximum and exceeds the corresponding threshold, the probability that the employer wins is given by*

$$\begin{aligned} &P_n(\text{Win}; \{\xi_n\}) \\ &= \frac{1}{n} + \left[ \frac{\xi_2^{n-1}}{(n-1) \cdot 1} - \frac{\bar{H}_1 \xi_2^n}{n} \right] + \left[ \frac{\xi_3^{n-2}}{(n-2) \cdot 2} + \frac{\xi_3^{n-1}}{(n-1) \cdot 1} - \frac{\bar{H}_2 \xi_3^n}{n} \right] \\ &\quad + \dots + \left[ \frac{\xi_n}{1 \cdot (n-1)} + \frac{\xi_n^2}{2 \cdot (n-2)} + \dots + \frac{\xi_n^{n-1}}{(n-1) \cdot 1} - \frac{\bar{H}_{n-1} \xi_n^n}{n} \right]. \end{aligned}$$

where  $\bar{H}_n = 1 + H_n$  and  $H_n = 1 + 1/2 + \dots + 1/n$  is the  $n$ -th harmonic number.

Using Theorem 2 we can find the employer's probability of winning if he uses any set of non-decreasing thresholds. In particular, in the simulation study of Section 3, for  $n = 5$ , we simply estimated  $P\{\text{Win}\}$  based on  $10^6$  plays of the game for the optimal strategy as well as for two alternative strategies. We can now compute the exact value of  $P\{\text{Win}\}$  for each strategy: It is .6391947 for optimal strategy, .6370689 for alternative 1, and .5923543 for alternative 2. Thus, we exhibit that indeed the alternative strategies are sub-optimal,

## 5 Conclusion

We reiterate that there is no need to restrict the candidates scores to have the uniform  $(0, 1)$  distribution. The results of this paper will continue to hold if the scores are drawn from any known continuous distribution function  $F$  with only one small change: Simply replace the thresholds of the employer's optimal strategy by the corresponding percentiles of  $F$ . The employer's maximal probability of winning will remain unaltered.

Also we can solve the problem of hiring the candidate with the minimum score (say, on some undesirable negative trait) among all  $n$  candidates by simply defining  $Y = 1 - X$  and hiring the candidate with the maximum  $Y$ -score. Likewise, we can solve the problem of hiring the candidate with a score closest to any specific percentile of  $F$ .

Whereas in the usual Secretary Problem (with only relative ranks available) the limiting (as  $n \rightarrow \infty$ ) winning probability is  $1/e = .368$ , in our score-based secretary problem, we conjecture that the limiting winning probability is about .580. The details are left to the reader.

## References

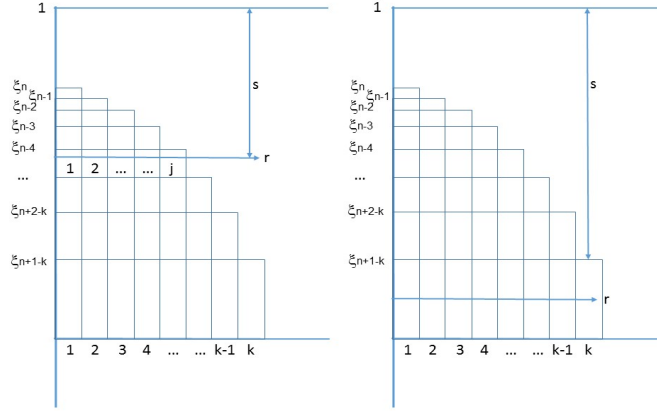
1. Bittinger, M.L. Ellenboggenn, D.J. and Surgent, S.J.: *Calculus and Its Applications*, 11th Ed. Pearson: Upper Saddle River, NJ. (2016)
2. Bruss, F.T.: A unified Approach to a Class of Best Choice Problems with an Unknown Number of Options. *Annals of Probability*. 12(3), 882-891 (1984)
3. Ferguson, T.S.: (1989). Who solved the secretary problem? *Statistical Science* 4(3), 282-296 (1989)
4. Flood, M.R.: A letter written in 1958. In: *The Martin Gardner Papers at Stanford University Archives*, series 1, box 5, folder 19.
5. Gardner, M.: *New Mathematical Diversions from Scientific American*: Chapter 3, Problem 3. Simon and Schuster, NY (1966)
6. Hogg, R.V., McKean, J., and Craig, A.T.: *Introduction to Mathematical Statistics*, 7th Ed. Pearson: Upper Saddle River, NJ. (2013)
7. Ross, S.M.: *Stochastic Processes*, 2nd Ed. John Wiley & Sons, New York, (1996)

## APPENDIX

**Proof of Theorem 2.** Suppose that the employer hires Candidate  $k$ , where  $k = 1, 2, \dots, n$ . Recall that the employer hires Candidate 1 iff  $X_1 > \xi_1$ . Otherwise, he hires Candidate  $k \geq 2$  provided he has not already hired anyone earlier and  $X_k = s$  exceeds both the previous maximum  $r = \max\{X_1, \dots, X_{k-1}\}$  and the threshold value  $\xi_{n+1-k}$  for Candidate  $k$ . In order to express the probability density function of  $r$ , subject to the fact that no candidate among Candidates  $1, 2, \dots, k-1$  has been hired, we must account for which candidate possibly achieved the record score  $r$  among Candidates  $1, 2, \dots, k-1$ . Therefore, we split the range  $(0, \xi_n)$  of  $r$  into sub-intervals

$$(\xi_{n-1}, \xi_n), (\xi_{n-2}, \xi_{n-1}), \dots, (\xi_{n+1-k}, \xi_{n+2-k}), (0, \xi_{n+1-k}) ;$$

and note that the number of possible candidates who might have achieved  $r$  is respectively  $1, 2, \dots, k-1, k-1$  in these intervals. See Fig. 6.



**Fig. 6.** If Candidate  $k \geq 2$  with score  $X_k = s$  is hired, then what is the likely score  $r = \max\{X_1, \dots, X_{k-1}\}$  of the previous leading candidate, and which candidate is she?

Having hired Candidate  $k$ , the employer wins (or hires the best candidate) only if all future Candidates also score below Candidate  $k$ . Adding up the joint probability that the employer hires Candidate  $k$  and he wins, we obtain

$$\begin{aligned} P_n(\text{Win}) &= \sum_{k=1}^n P_n(\text{Hire } k \text{ and Win}) \\ &= \int_{\xi_n}^1 s^{n-1} ds + \sum_{k=2}^{n-1} \left\{ \sum_{j=1}^{k-1} \int_{\xi_{n-j}}^{\xi_{n+1-j}} j r^{k-2} \int_r^1 s^{n-k} ds dr \right\} \end{aligned}$$

$$\begin{aligned}
& + \int_0^{\xi_{n+1-k}} (k-1) r^{k-2} \int_{\xi_{n+1-k}}^1 s^{n-k} ds dr \Big\} \\
& + \sum_{j=1}^{n-1} \int_{\xi_{n-j}}^{\xi_{n+1-j}} j r^{n-2} \int_r^1 s^0 ds dr \\
& = \frac{1 - \xi_n^n}{n} + \sum_{k=2}^{n-1} \left\{ \sum_{j=1}^{k-1} \int_{\xi_{n-j}}^{\xi_{n+1-j}} j \frac{r^{k-2} - r^{n-1}}{n+1-k} dr \right. \\
& \quad \left. + \int_0^{\xi_{n+1-k}} (k-1) r^{k-2} \frac{1 - \xi_{n+1-k}^{n+1-k}}{n+1-k} dr \right\} \\
& \quad + \sum_{j=1}^{n-1} \int_{\xi_{n-j}}^{\xi_{n+1-j}} j (r^{n-2} - r^{n-1}) dr \\
& = \frac{1 - \xi_n^n}{n} + \sum_{k=2}^n \frac{1}{n+1-k} \left\{ \sum_{j=1}^{k-1} j \left( \frac{\xi_{n-j}^{k-1} - \xi_{n+1-j}^{k-1}}{k-1} - \frac{\xi_{n-j}^n - \xi_{n+1-j}^n}{n} \right) \right. \\
& \quad \left. + (\xi_{n+1-k}^{k-1} - \xi_{n+1-k}^n) \right\} \\
& = \frac{1}{n} [1 - \xi_n^n - \xi_{n-1}^n - \dots - \xi_2^n - \xi_1^n] \\
& \quad + \frac{1}{n-1} \left( \xi_n - \frac{\xi_n^n}{n} \right) + \frac{1}{n-2} \left[ \left( \frac{\xi_n^2}{2} - \frac{\xi_n^n}{n} \right) + \left( \frac{\xi_{n-1}^2}{2} - \frac{\xi_{n-1}^n}{n} \right) \right] \\
& \quad + \frac{1}{n-3} \left[ \left( \frac{\xi_n^3}{3} - \frac{\xi_n^n}{n} \right) + \left( \frac{\xi_{n-1}^3}{3} - \frac{\xi_{n-1}^n}{n} \right) + \left( \frac{\xi_{n-2}^3}{3} - \frac{\xi_{n-2}^n}{n} \right) \right] \\
& \quad + \dots \\
& \quad + \frac{1}{1} \left[ \left( \frac{\xi_n^{n-1}}{n-1} - \frac{\xi_n^n}{n} \right) + \left( \frac{\xi_{n-1}^{n-1}}{n-1} - \frac{\xi_{n-1}^n}{n} \right) + \dots + \left( \frac{\xi_2^{n-1}}{n-1} - \frac{\xi_2^n}{n} \right) \right] \\
& = \frac{1}{n} [1 - \xi_n^n - \xi_{n-1}^n - \dots - \xi_2^n] \\
& \quad + \left[ \frac{\xi_n}{1 \cdot (n-1)} + \frac{\xi_n^2 + \xi_{n-1}^2}{2 \cdot (n-2)} + \dots + \frac{\xi_n^{n-1} + \xi_{n-1}^{n-1} + \dots + \xi_2^{n-1}}{(n-1) \cdot 1} \right] \\
& \quad - \frac{\xi_n^n}{n} H_{n-1} - \frac{\xi_{n-1}^n}{n} H_{n-2} - \dots - \frac{\xi_2^n}{n} H_1,
\end{aligned}$$

whence the theorem follows.

Q.E.D.